Musterlösung

1. **a)** Richtig, da es sich hier um die Anzahl Erfolge (Kopf) bei $n = 10$ unabhängigen Münzwürfen handelt und P(Kopf) $= 0.5$.

b) Falsch, die hypergeometrische Verteilung wäre passend.

c) Falsch. Als Faustregel gilt, dass die Normalapproximation gut ist, falls $n\pi > 5$ und $n(1-\pi) > 5$. Das ist hier nicht der Fall.

d) Richtig, da für $n = 1$ die Binomialverteilung der Bernoulliverteilung entspricht und jede Zahl mit Wahrscheinlichkeit $p = 1/100$ oben liegt.

2. **a)** Falsch, da $E[X] = np = 0.2 \times 8 = 1.60$.

b) Richtig, da $Var(X) = np(1-p) = 1.28$.

c) Falsch. Die Erfolgswahrscheinlichkeit ist $p = 0.2$ und nicht 0.8.

d) Richtig. $P[X <= 2] = P[X = 0] + P[X = 1] + P[X = 2] = 0.80$

3. **a)** Richtig. Die Macht des einseitigen Tests ist also viel grösser als die des zweiseitigen Tests.

b) Richtig.

c) Falsch, $T \sim Bin(n = 15, p = 0.3)$.

d) Falsch. Das Signifikanzniveau beschränkt den Fehler 1. Art.

4. **a)** Richtig, $T \sim Bin(n = 10, p = 0.4)$ und der Verwerfungsbereich ist $K = [0, c]$. Aus $P(X = 0) = 0.01$, $P(X = 1) = 0.04$ und $P(X = 2) = 0.12$ folgt $c = 1$.

b) Falsch, der Verwerfungsbereich wird dann tendenziell weniger Elemente enthalten.

c) Richtig.

d) Falsch, der Test-Entscheidung beruht auf dem Widerspruchs-Prinzip: Hypothesen können nur falsifiziert und nicht verifiziert werden.

5. **a)** Falsch, die Macht ist 1 - P(Fehler 2. Art).

b) Richtig, die Macht berechnet sich als
$P_{H_A}(X \in K) = P_{p=0.25}(X = 0) + P_{p=0.25}(X = 1) = 0.1 + 0.27 = 0.37$.

c) Falsch, die Macht für die Alternative $H_A:\ p = 0.35$ ist 0.17, also kleiner.

d) Richtig, da Macht = P (Verwerfen von H_0 falls H_A stimmt).

6. **a)** Richtig, H_0 wird verworfen, falls der p-Wert $= 0.037 \leq 0.05 = \alpha$.

b) Richtig, der p-Wert ist $P_{p=0.5}(X \leq 0) = P_{p=0.5}(X = 0) = 0.5^{10} \approx 0.001 < 0.01$.

c) Richtig, das Vertrauensintervall ist gegeben durch

$$I \approx \frac{17}{30} \pm 1.96\sqrt{\frac{17}{30}(1 - \frac{17}{30})\frac{1}{30}} = [0.39, 0.74]$$

d) Falsch, das 95%-Vertrauensintervall enthält den wahren Parameter mit einer Wahrscheinlichkeit von 95%.

7. **a)** Richtig.

b) Falsch.

c) Falsch. Falls eine Paarung vorliegt, hat der gepaarte Test tendenziell eine grössere Macht.

d) Falsch.

8. **a)** Richtig, man möchte im Sinne des Besitzers statistisch nachweisen, dass die Temperaturen im Jahr 2012 mehr als 2°C höher waren als im Jahr 2013, also dass die Differenz der Temperaturen im Jahr 2012 und 2013 grösser als 2 ist. Die einseitige Alternativhypothese von der Form $H_A : \mu_D > 2$ hat dabei die grösste Macht.

b) Richtig.

c) Falsch, die Verteilung von T unter H_0 ist $T \sim t_{n-1} = t_4$.

d) Falsch. Der richtige Verwerfungsbereich wäre $[-\infty, -2.132]$ (also mit Minuszeichen).

9. **a)** Richtig. Angenommen $T\ t_{11}$. Der p-Wert berechnet sich dann als $p = P(T < -1.3 \cup T > 1.3) = 2 * P(T > 1.3)$. Aus der Tabelle für die t-Verteilung können wir abschätzen, dass $P(T > 1.3) = 1 - P(T \leq 1.3) \approx 1 - 0.9 = 0.1$. Also ist $p \approx 0.2$.

b) Richtig.

c) Falsch. $[2.5 \pm 2.78 * 0.64/sqrt(5)]$.

d) Falsch.

10. **a)** Richtig. Ein grösseres Intervall hat auch eine grössere Wa. den wahren Parameter zu überdecken. Ein 100%-Vertrauensintervall ist z.B. $[-\infty; \infty]$.

b) Richtig.

c) Falsch.

d) Richtig. Mit grosser Wa. (mind. 95%) ist der Wert von μ_d im 95%-Vertrauensintervall. Für alle Werte in diesem Intervall ist aber der Absoutbetrag kleiner als 3.0. Also besteht mit grosser Wa. kein relevanter Unterschied.

11. **a)** Falsch, Oskar und Max haben 35 Messungen gemacht (Anzahl degrees of freedom (32) plus Anzahl betas (3))

b) Falsch. Der Verwerfungsbereich ist $K = (-\infty, -t_{32,0.975}] \cup [t_{32,0.975}, \infty) = (-\infty, -2.037] \cup [2.037, \infty)$ und $t = \frac{1.987}{2.572} = 0.773$ ist nicht in K.

c) Richtig. Die Schätzung berechnet sich als se(b)$\cdot$ t-value(b) $= 2.326 \cdot (-0.82) = -1.91$.

d) Falsch. Das exakte zweiseitige 99%-Vertrauensintervall für β_2 ist:
$[\hat{\beta}_2 - t_{32,0.995} \cdot se(\hat{\beta}_2), \hat{\beta}_2 + t_{32,0.995} \cdot se(\hat{\beta}_2)] = [5.340 - 2.738 \cdot 0.492, 5.340 + 2.738 \cdot 0.492] = [3.99, 6.69]$

12. **a)** Falsch, gemäss dem Modell ist der Ball aus etwa 44m gestartet.

b) Falsch, das Residuum beträgt 0.5 Meter.

c) Falsch. Es ist das Vertrauensintervall gegeben. Um die Frage beantworten zu können, bräuchten wir das Vorhersageintervall.

d) Richtig, da zusätzlich noch die Streuung durch den Zufallsfehler hinzukommt.

13. **a)** Falsch, da das Modell linear in den Koeffizienten ist.

b) Richtig.

c) Richtig.

d) Falsch.

14. **a)** Falsch. $P(A \cup B) = P(A) + P(B) - P(A \cap B) = 0.7 + 0.5 - P(A \cap B) = 1.2 - P(A \cap B)$. Zudem muss eine Wa. immer höchstens eins sein.

b) Richtig. Die odds für A sind $P(A)/(1 - P(A)) = 0.6/0.4 = 1.5$.

c) Falsch. Für eine stetige Zufallsvariable gilt $P(X = x) = 0$ für alle x im Wertebereich von X.

d) Falsch. Nur wenn A und B sind unabhängig sind, gilt im Allgemeinen $P(A) * P(B) = 0.2 * 0.5 = 0.1 = P(A \cap B)$.

15. **a)** Richtig, da $E[X - 2Y + 2] = E[X] - 2E[Y] + 2 = 2 - 6 + 2 = -2$

b) Falsch, die richtige Antwort ist: $Var(X + 3Y - 5) = Var(X) + 9Var(Y) = 16 + 9 * 3 = 43$.

c) Richtig. $P(B) = P(B|A = 1)P(A = 1) + P(B|A = 2)P(A = 2) = 0.2 * 0.1 + 0.1 * 0.9 = 0.11$ und damit $P(A = 1 \mid B) = P(B|A = 1)P(A = 1)/P(B) = 0.2 * 0.1/0.11 = 0.18$.

d) Falsch. Nach dem Satz von Bayes: $P(A|B) = P(B|A)P(A)/P(B)$.

16. a) Richtig. Die dicke Linie im Rechteck zeigt den Median an. Diese ist etwa bei 1.

b) Falsch. Nur 25% der Daten befinden sich oberhalb des Rechtecks, also sind weniger als 25% der Daten grösser als 2.

c) Falsch. Die Daten sind rechtsschief, somit ist der Mittelwert grösser als der Median.

d) Falsch. Oberhalb und unterhalb des Medians sind genau gleich viele Datenpunkte.

17. a) Falsch, die Fläche unter der Kurve ist 4. Für eine Dichte gilt jedoch, dass die Fläche unter der Kurve 1 ist.

b) Richtig. Sei S_{100} der Gewinn nach 100 Spielen.
Aus dem Zentralen Grenzwertsatz folgt:

$$S_{100} \sim \mathcal{N}(100\,(-1), 3^2\,100).$$

Daher

$$P(S_{100} \leq -40) = P\left(\frac{(S_{100} + 100)}{\sqrt{900}} \leq \frac{(-40 + 100)}{\sqrt{900}}\right) = \Phi(2) = 0.9772$$

und

$$P(S_{100} \leq -160) = \Phi\left(\frac{-160 + 100)}{\sqrt{900}}\right) = \Phi(-2) = 0.0228$$

Somit ist $P(-160 < S_{100} < -40) = P(S_{100} < -40) - P(S_{100} < -160) = 0.95$.

c) Falsch wegen Wurzel-n-Gesetz.

d) Falsch.

18. a) Falsch. Es gilt $E[Y] = E[2X] = 2 \cdot E[X] = 2 \cdot 3$ und $\mathrm{Var}(Y) = \mathrm{Var}(2X) = 2^2\,\mathrm{Var}(X) = 2^2 \cdot 5$.

b) Richtig. Die Wahrscheinlichkeit $P(X \leq 2)$ ist gerade die Fläche unter der Dichtefunktion bis $x = 2$. Deshalb gilt: $P(X \leq 2) = 1 * 0.1 + 2 * 0.3 = 0.7$.

c) Falsch. Da sowohl Y als auch Z lineare Funktionen von X sind mit positiver Steigung, gilt: Corr$(X, Y) =$ Corr$(X, Z) = 1$.

d) Falsch. Die Poissonverteilung beschreibt die Anzahl Ereignisse während eines Zeitintervalls.

19. a) Falsch, $P(a \leq X \leq b) = F(b) - F(a)$

b) Richtig.

c) Richtig.

d) Richtig.

20. a) Richtig. Wenn $X \sim \mathrm{Poisson}(\lambda_X)$ und $Y \sim \mathrm{Poisson}(\lambda_y)$ unabhängig sind, dann ist $X + Y \sim \mathrm{Poisson}(\lambda_X + \lambda_y)$.

b) Falsch.

c) Richtig. $\mathrm{Var}(2X - Y) = 4\mathrm{Var}(X) + \mathrm{Var}(Y) = 5\lambda$.

d) Richtig. Für grosse n und kleine p gilt: $\mathrm{Binomial}(n, p) \approx \mathrm{Poisson}(\lambda)$ mit $\lambda = np$.